mml-Linear Algebra

Linear Algebra

Systems of Linear Equations

For a systematic approach to solving systems of linear equations, we collect

- coefficients a_{ij} into **vectors**
- and vectors into **matrix**

$$\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} x_2 + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \qquad \Longleftrightarrow \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

Matrix

> *With $m, n \in N$ a real-valued (m, n) matrix A is an $m \cdot n$-tuple of elements a_{ij}, $i = 1, 2, ..., 3$, $j = 1, 2, ..., 3$ which is ordered according to a rectangular scheme consisting of m rows and n columns*

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad a_{ij} \in \mathbb{R}.$$

Addition and Multiplication

Addition

The tow matrices should be in the same shape.

$$\boldsymbol{A}+\boldsymbol{B}:=\begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n}.$$

Multiplication

The matrix of B should have the same rows as the columns of the matrix of A, multiply the elements of the ith row of A with the jth column of B and sum them up.

For matrices $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times k}$, the elements c_{ij} of the product $\boldsymbol{C} = \boldsymbol{AB} \in \mathbb{R}^{m\times k}$ are computed as

$$c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}, \qquad i = 1, \ldots, m, \quad j = 1, \ldots, k. \tag{2.13}$$

The multiplication of matrices is actually *dot production of corresponding row and column*, and use $A \cdot B$ to denote the operation.

> 🗣 $\mathbf{AB} \neq \mathbf{BA}$: more likely

Identify Matrix $\mathbf{I}_n$

Define a $n \times n$ - matrix contains 1 on the whole diagonal and 0 everywhere else.

$$\boldsymbol{I}_n := \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}$$

Some properties of matrix

Associativity

$\forall A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times p}, C \in \mathbb{R}^{p\times q} : (\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$

Distributivity

$\forall \mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}, \mathbf{C} \in \mathbb{R}^{n \times p} : (\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$

Multiplication with the $\mathbf{I}_n$

$\forall \mathbf{A} \in \mathbb{R}^{m \times n} : \mathbf{I}_m \mathbf{A} = \mathbf{A} \mathbf{I}_n$

Inverse Matrices

For square matrices, $\mathbf{AB} = \mathbf{I}_n = \mathbf{BA}$, then $\mathbf{B}$ is the inverse of the matrix $\mathbf{A}$, denoted by $\mathbf{A}^{-1}$.

> Not every matrix $\mathbf{A}$ has a inverse.

If the inverse of a matrix $\mathbf{A}$ exists, then $\mathbf{A}$ is called *regular/invertible/nonsingular*, and it's unique. Otherwise, $\mathbf{A}$ is called *singular/noninvertible.*

For a 2×2 - matrix, its inverse is, if and only if $a_{11}a_{22} - a_{12}a_{21} \neq 0$,

$$\boldsymbol{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Transpose Matrix

For tow matrices $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{B} \in \mathbb{R}^{n \times m}$ with $b_{ij} = a_{ji}$, then $\mathbf{B}$ is the *transpose* of $\mathbf{A}$, denoted by $\mathbf{B} = \mathbf{A}^T$.

More properties…

$$\begin{aligned} \boldsymbol{A}\boldsymbol{A}^{-1} &= \boldsymbol{I} = \boldsymbol{A}^{-1}\boldsymbol{A} \\ (\boldsymbol{A}\boldsymbol{B})^{-1} &= \boldsymbol{B}^{-1}\boldsymbol{A}^{-1} \\ (\boldsymbol{A}+\boldsymbol{B})^{-1} &\neq \boldsymbol{A}^{-1}+\boldsymbol{B}^{-1} \end{aligned} \qquad \begin{aligned} (\boldsymbol{A}^\top)^\top &= \boldsymbol{A} \\ (\boldsymbol{A}+\boldsymbol{B})^\top &= \boldsymbol{A}^\top+\boldsymbol{B}^\top \\ (\boldsymbol{A}\boldsymbol{B})^\top &= \boldsymbol{B}^\top\boldsymbol{A}^\top \end{aligned}$$

Symmetric Matrix

If $\mathbf{A} = \mathbf{A}^T$, then $\mathbf{A}$ is symmetric.

<u>Only square matrices can be symmetric.</u>

If $\mathbf{A}$ is invertible, then so is $\mathbf{A}^T$: $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1} := \mathbf{A}^{-T}$.

The sum of tow symmetric matrices is always symmetric, the product, however, is generally not.

Multiplication by a scalar

When a scalar $\lambda \in \mathbb{R}$ multiply to a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, the λ times with each element in the matrix, the result is also a matrix.

Properties…

- *Associativity:*
  $(\lambda\psi)\boldsymbol{C} = \lambda(\psi\boldsymbol{C}), \quad \boldsymbol{C} \in \mathbb{R}^{m\times n}$
- $\lambda(\boldsymbol{BC}) = (\lambda\boldsymbol{B})\boldsymbol{C} = \boldsymbol{B}(\lambda\boldsymbol{C}) = (\boldsymbol{BC})\lambda, \quad \boldsymbol{B} \in \mathbb{R}^{m\times n}, \boldsymbol{C} \in \mathbb{R}^{n\times k}$.
  Note that this allows us to move scalar values around.
- $(\lambda\boldsymbol{C})^\top = \boldsymbol{C}^\top\lambda^\top = \boldsymbol{C}^\top\lambda = \lambda\boldsymbol{C}^\top$ since $\lambda = \lambda^\top$ for all $\lambda \in \mathbb{R}$.
- *Distributivity:*
  $(\lambda+\psi)\boldsymbol{C} = \lambda\boldsymbol{C} + \psi\boldsymbol{C}, \quad \boldsymbol{C} \in \mathbb{R}^{m\times n}$
  $\lambda(\boldsymbol{B}+\boldsymbol{C}) = \lambda\boldsymbol{B} + \lambda\boldsymbol{C}, \quad \boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}^{m\times n}$

Compact Representations of Linear Equations

Represent linear equations in matrix form as $\mathbf{A}x = b$, with $\mathbf{A}$ represents the coefficients, $\mathbf{x}$ represents the xs.

Solving System of Linear Equations

Particular and General Solution

The Minus-1 Trick

for reading out the solutions of a homogeneous system of linear equations $Ax = 0$, where $A \in \mathbb{R}^{k\times n}$, by extending the matrix(RREF) to $n \times n$-matrix with $n - k$ rows of

$$\begin{bmatrix} 0 & \cdots & 0 & -1 & 0 & \cdots & 0 \end{bmatrix}$$

> Example of Minus-1 Trick

Example 2.8 (Minus-1 Trick)
Let us revisit the matrix in (2.49), which is already in reduced REF:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \end{bmatrix}. \tag{2.53}$$

We now augment this matrix to a 5×5 matrix by adding rows of the form (2.52) at the places where the pivots on the diagonal are missing and obtain

$$\tilde{\boldsymbol{A}} = \begin{bmatrix} 1 & 3 & 0 & 0 & 3 \\ 0 & -\mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & -\mathbf{1} \end{bmatrix}. \tag{2.54}$$

From this form, we can immediately read out the solutions of $\boldsymbol{Ax} = \mathbf{0}$ by taking the columns of $\tilde{\boldsymbol{A}}$, which contain -1 on the diagonal:

$$\left\{ \boldsymbol{x} \in \mathbb{R}^5 : \boldsymbol{x} = \lambda_1 \begin{bmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 3 \\ 0 \\ 9 \\ -4 \\ -1 \end{bmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R} \right\}, \tag{2.55}$$

which is identical to the solution in (2.50) that we obtained by “insight”.

Elementary Transformation

Transform the equation system into a simpler form, using the Gaussian Elimination laws.

- Exchange of two equations (rows in the matrix representing the system of equations)
- Multiplication of an equation (row) with a constant $\lambda \in \mathbb{R}\backslash\{0\}$
- Addition of two equations (rows)

The Gaussian elimination can change augmented matrix into the ***reduced row-echelon form(RREF)***.

An example of a REF of a $4 \times 5-$matrix.

$$\begin{bmatrix} 1 & a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & 2 & a_4 & a_5 \\ 0 & 0 & 0 & 1 & a_6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Pivots: the leading coefficient of a row, or, 1st nonzero number from left

Basic and free variables:

- basic variables are those corresponding to the pivots, x_1, x_3, x_4 in the above example.
- the others are free variables, x_2, x_5 in the above example.

Reduced Row Echelon Form(RREF): a REF where every pivot is 1 and is the only nonzero entry in its column. An example of a RREF $3 \times 5-$matrix.

$$\begin{bmatrix} 1 & 0 & a_1 & 0 & b_1 \\ 0 & 1 & a_2 & 0 & b_2 \\ 0 & 0 & 0 & 1 & b_3 \end{bmatrix}$$

The REF helps find the *particular solution* while the *RREF* helps find the *general solution* directly.

Calculating the Inverse

To compute the inverse A^{-1} of $A \in \mathbb{R}^{n \times n}$, we need a matrix X that satisfies $AX = I_n$, then $X = A^{-1}$.

We use the augmented matrix $[A|I_n] \rightarrow [I_n|A^{-1}]$ to reduced row-echelon form, the the inverse is on the right-hand side of the equation system.

Moore-Penrose pseudo-inverse

For square and invertible matrices, we can use the inverse to solve the equation of $Ax = b$ with the result as $x = A^{-1}b$.

For more usual cases, we can use the transformation

$$Ax = b \iff A^T Ax = A^T b \iff x = (A^T A)^{-1} A^T b$$

and use the *Moore-Penrose pseudo-inverse* $(A^T A)^{-1} A^T$ to determine the solution.

> 🗣️ Due to large computation of matrix production and inverse, it’s not recommended.

Vector Spaces

Groups

A set of elements and an operation defined on these elements that keeps some structure of the set intact.

> For a set ς and an operation $\otimes$, $G := (\varsigma, \otimes)$ is a group if:
>
> 1. *Closure* of $\mathcal{G}$ under $\otimes$: $\forall x, y \in \mathcal{G} : x \otimes y \in \mathcal{G}$
> 2. *Associativity:* $\forall x, y, z \in \mathcal{G} : (x \otimes y) \otimes z = x \otimes (y \otimes z)$
> 3. *Neutral element:* $\exists e \in \mathcal{G} \forall x \in \mathcal{G} : x \otimes e = x$ and $e \otimes x = x$
> 4. *Inverse element:* $\forall x \in \mathcal{G} \exists y \in \mathcal{G} : x \otimes y = e$ and $y \otimes x = e$, where e is the neutral element. We often write x^{-1} to denote the inverse element of x.

Abelian group / commutative:

If additionally $\forall x, y \in \varsigma : x \otimes y = y \otimes x$, then $G = (\varsigma, \otimes)$ is Abelian group.

> 🗣️ $(\mathbb{R}^{m \times n}, +)$, the set of $m \times n-$matrix is Abelian, with component-wise addition.

General linear group:

$(\mathbb{R}^{n \times n}, \cdot)$ is the *general linear group*, denoted by $GL(n, \mathbb{R})$, when the the matrices are regular, or invertible and with the respect to matrix multiplication as (2.13).

Since the matrix multiplication is not commutative, the GL-group is not Abelian.

Vector Spaces

Inner operation → +

Outer operation → •, the multiplication of a vector by a scalar

> A real-valued vector space $V = (\gamma, +, \cdot)$ is a set γ with 2 operations

$$+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V} \tag{2.62}$$

$$\cdot : \mathbb{R} \times \mathcal{V} \to \mathcal{V} \tag{2.63}$$

where

1. $(\mathcal{V}, +)$ is an Abelian group
2. Distributivity:
    1. $\forall \lambda \in \mathbb{R}, \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V} : \lambda \cdot (\boldsymbol{x} + \boldsymbol{y}) = \lambda \cdot \boldsymbol{x} + \lambda \cdot \boldsymbol{y}$
    2. $\forall \lambda, \psi \in \mathbb{R}, \boldsymbol{x} \in \mathcal{V} : (\lambda + \psi) \cdot \boldsymbol{x} = \lambda \cdot \boldsymbol{x} + \psi \cdot \boldsymbol{x}$
3. Associativity (outer operation): $\forall \lambda, \psi \in \mathbb{R}, \boldsymbol{x} \in \mathcal{V} : \lambda \cdot (\psi \cdot \boldsymbol{x}) = (\lambda\psi) \cdot \boldsymbol{x}$
4. Neutral element with respect to the outer operation: $\forall \boldsymbol{x} \in \mathcal{V} : 1 \cdot \boldsymbol{x} = \boldsymbol{x}$

- The elements $\mathbf{x} \in \gamma$ are *vectors*
- The neutral element of $(\gamma, +)$ is *zero vector* $\mathbf{0} = [0, \ldots, 0]^{\top}$
- Inner operation + is *vector addition*
- The elements $\lambda \in \mathbb{R}$ are *scalar*, the outer operation • is *multiplication by scalars*

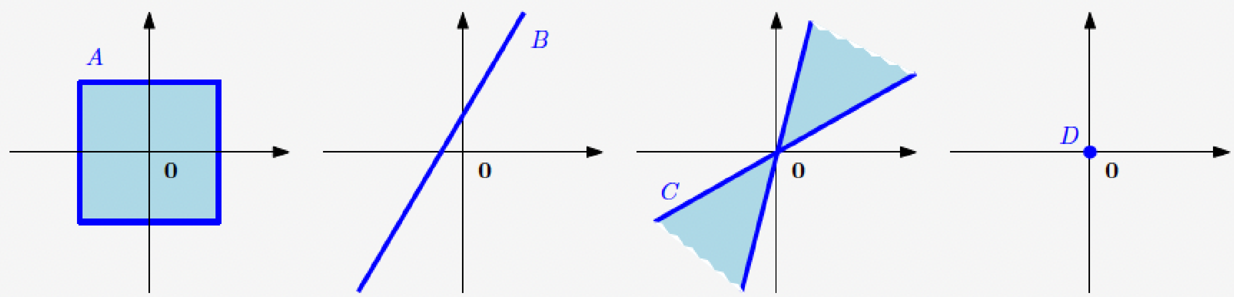
Vector Subspaces

> vector subspace
>
> **Definition 2.10** (Vector Subspace). Let $V = (\mathcal{V}, +, \cdot)$ be a vector space and $\mathcal{U} \subseteq \mathcal{V}, \mathcal{U} \neq \emptyset$. Then $U = (\mathcal{U}, +, \cdot)$ is called *vector subspace* of V (or *linear subspace*) if U is a vector space with the vector space operations $+$ and $\cdot$ restricted to $\mathcal{U} \times \mathcal{U}$ and $\mathbb{R} \times \mathcal{U}$. We write $U \subseteq V$ to denote a subspace U of V.

To determine whether $(U, +, \cdot)$ is a subspace

1. $\mathcal{U} \neq \emptyset$, in particular: $\mathbf{0} \in \mathcal{U}$
2. Closure of U:
    a. With respect to the outer operation: $\forall \lambda \in \mathbb{R} \, \forall \boldsymbol{x} \in \mathcal{U} : \lambda \boldsymbol{x} \in \mathcal{U}$.
    b. With respect to the inner operation: $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{U} : \boldsymbol{x} + \boldsymbol{y} \in \mathcal{U}$.

- For every vector space V, the trivial subspaces are V itself and $\{\mathbf{0}\}$.
- Only example D in Figure 2.1 is a subspace of $\mathbb{R}^2$ (with the usual inner/outer operations). In A and C, the closure property is violated; B does not contain $\mathbf{0}$.
- The solution set of a homogeneous system of linear equations $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ with n unknowns $\boldsymbol{x} = [x_1, \ldots, x_n]^\top$ is a subspace of $\mathbb{R}^n$.
- The solution of an inhomogeneous system of linear equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{b} \neq \mathbf{0}$ is not a subspace of $\mathbb{R}^n$.
- The intersection of arbitrarily many subspaces is a subspace itself.



Linear Independence

> ### Linear Combination
>
> **Definition 2.11** (Linear Combination)**.** Consider a vector space V and a finite number of vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in V$. Then, every $\boldsymbol{v} \in V$ of the form
>
> $$\boldsymbol{v} = \lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_k \boldsymbol{x}_k = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i \in V \tag{2.65}$$
>
> with $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ is a *linear combination* of the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$.

$0-$vector is always a linear combination, since $0 = \sum_{i=1}^{k} 0x_i$ is always true.

> ### Linear Independence
>
> If there's at least one coefficient isn't 0, then it's *linearly dependent.*

> **Definition 2.12** (Linear (In)dependence)**.** Let us consider a vector space V with $k \in \mathbb{N}$ and $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in V$. If there is a non-trivial linear combination, such that $\mathbf{0} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i$ with at least one $\lambda_i \neq 0$, the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ are *linearly dependent*. If only the trivial solution exists, i.e., $\lambda_1 = \ldots = \lambda_k = 0$ the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ are *linearly independent*.

Useful properties to check linearly independency

- k vectors are either *linearly dependent* or *linearly independent*, no other options.
- If at least one of the vectors is $\mathbf{0}$ then they're *linearly dependent*. Same if there're 2 identical vectors.
- Using Gaussian Elimination to convert matrix to (Reduced) REF, and
  - if and only if all columns are pivot columns, they are *linear independent*
  - if there's one non-pivot column, they're *linearly dependent*

Basis and Rank

Basis

> #### Generating Set and Span
>
> **Definition 2.13** (Generating Set and Span)**.** Consider a vector space $V = (\mathcal{V}, +, \cdot)$ and set of vectors $\mathcal{A} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} \subseteq \mathcal{V}$. If every vector $\boldsymbol{v} \in \mathcal{V}$ can be expressed as a linear combination of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$, $\mathcal{A}$ is called a *generating set* of V. The set of all linear combinations of vectors in $\mathcal{A}$ is called the *span* of $\mathcal{A}$. If $\mathcal{A}$ spans the vector space V, we write $V = \text{span}[\mathcal{A}]$ or $V = \text{span}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k]$.

Generating sets are sets of vectors that span vector (sub)spaces, every vector can be represented as a linear combination of the vectors in the generating set.

> the smallest generating set that spans a vector (sub)space

Definition 2.14 (Basis). Consider a vector space $V = (\mathcal{V}, +, \cdot)$ and $\mathcal{A} \subseteq \mathcal{V}$. A generating set $\mathcal{A}$ of V is called *minimal* if there exists no smaller set $\tilde{\mathcal{A}} \subsetneq \mathcal{A} \subseteq \mathcal{V}$ that spans V. Every linearly independent generating set of V is minimal and is called a *basis* of V.

A basis is a minimal generating set and a maximal linearly independent set of vectors.

Let $V = (\mathcal{V}, +, \cdot)$ be a vector space and $\mathcal{B} \subseteq \mathcal{V}, \mathcal{B} \neq \emptyset$. Then, the following statements are equivalent:

- $\mathcal{B}$ is a basis of V.
- $\mathcal{B}$ is a minimal generating set.
- $\mathcal{B}$ is a maximal linearly independent set of vectors in V, i.e., adding any other vector to this set will make it linearly dependent.
- Every vector $\boldsymbol{x} \in V$ is a linear combination of vectors from $\mathcal{B}$, and every linear combination is unique, i.e., with

$$\boldsymbol{x} = \sum_{i=1}^{k} \lambda_i \boldsymbol{b}_i = \sum_{i=1}^{k} \psi_i \boldsymbol{b}_i \tag{2.77}$$

and $\lambda_i, \psi_i \in \mathbb{R}$, $\boldsymbol{b}_i \in \mathcal{B}$ it follows that $\lambda_i = \psi_i,\ i = 1, \ldots, k$.

- Every vector space possesses a basis, but there's no unique basis.
- *Basis vectors*, the same number of elements of basis.

In finite-dimensional vector spaces V, the *dimension* of V is the number of basis vectors of V, write $dim(V)$.

- If $U \subseteq V$ is a subspace of V, then $dim(U) \leq dim(V)$ and $dim(U) = dim(V)$ if and only if $U = V$.

Determining a Basis

Determine a basis of a subspace:

Remark. The dimension of a vector space is not necessarily the number of elements in a vector.

Remark. A basis of a subspace $U = \text{span}[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m] \subseteq \mathbb{R}^n$ can be found by executing the following steps:

1. Write the spanning vectors as columns of a matrix $\boldsymbol{A}$
2. Determine the row-echelon form of $\boldsymbol{A}$.
3. The spanning vectors associated with the pivot columns are a basis of U.

> 🗣️ Make vectors to a matrix → Gaussian Elimination transfer to REF → the pivots columns are the basis of the subspace U

Rank

> The number of linearly independent columns of a matrix $A \in \mathbb{R}^{m \times n}$ equals to the number of linearly independent rows and is called the *rank*, denoted by $rk(\mathbf{A})$.

The rank is also the number of pivots of a Row-Echelon Form of the matrix A.

- $\text{rk}(\boldsymbol{A}) = \text{rk}(\boldsymbol{A}^\top)$, i.e., the column rank equals the row rank.
- The columns of $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ span a subspace $U \subseteq \mathbb{R}^m$ with $\dim(U) = \text{rk}(\boldsymbol{A})$. Later we will call this subspace the *image* or *range*. A basis of U can be found by applying Gaussian elimination to $\boldsymbol{A}$ to identify the pivot columns.
- The rows of $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ span a subspace $W \subseteq \mathbb{R}^n$ with $\dim(W) = \text{rk}(\boldsymbol{A})$. A basis of W can be found by applying Gaussian elimination to $\boldsymbol{A}^\top$.
- For all $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ it holds that $\boldsymbol{A}$ is regular (invertible) if and only if $\text{rk}(\boldsymbol{A}) = n$.
- For all $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and all $\boldsymbol{b} \in \mathbb{R}^m$ it holds that the linear equation system $\boldsymbol{Ax} = \boldsymbol{b}$ can be solved if and only if $\text{rk}(\boldsymbol{A}) = \text{rk}(\boldsymbol{A}|\boldsymbol{b})$, where $\boldsymbol{A}|\boldsymbol{b}$ denotes the augmented system.
- For $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ the subspace of solutions for $\boldsymbol{Ax} = \boldsymbol{0}$ possesses dimension $n - \text{rk}(\boldsymbol{A})$. Later, we will call this subspace the *kernel* or the *null space*.
- A matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ has *full rank* if its rank equals the largest possible rank for a matrix of the same dimensions. This means that the rank of a full-rank matrix is the lesser of the number of rows and columns, i.e., $\text{rk}(\boldsymbol{A}) = \min(m, n)$. A matrix is said to be *rank deficient* if it does not have full rank.

Linear Mappings

> *Vectors are objects that can be added together and multiplied by a scalar, and the resulting object is still a vector*
>
> **Definition 2.15** (Linear Mapping)**.** For vector spaces V, W, a mapping $\Phi : V \to W$ is called a *linear mapping* (or *vector space homomorphism/ linear transformation*) if
>
> $$\forall \boldsymbol{x}, \boldsymbol{y} \in V\ \forall \lambda, \psi \in \mathbb{R} : \Phi(\lambda\boldsymbol{x} + \psi\boldsymbol{y}) = \lambda\Phi(\boldsymbol{x}) + \psi\Phi(\boldsymbol{y})\,. \tag{2.87}$$

We can represent linear mappings as matrices.

Homomorphism

> A homomorphism is a map between two algebraic structures of the same type (that is of the same name), (here is tow vector space), that preserves

> the <u>operations</u> of the structures.

For example:

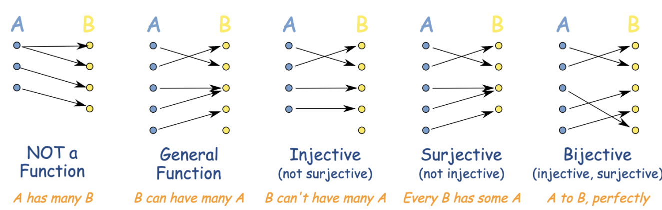
The mapping $\Phi : \mathbb{R}^2 \to \mathbb{C}$, $\Phi(\boldsymbol{x}) = x_1 + ix_2$, is a homomorphism:

$$\begin{aligned}\Phi\left(\begin{bmatrix}x_1\\x_2\end{bmatrix} + \begin{bmatrix}y_1\\y_2\end{bmatrix}\right) &= (x_1+y_1) + i(x_2+y_2) = x_1 + ix_2 + y_1 + iy_2\\ &= \Phi\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) + \Phi\left(\begin{bmatrix}y_1\\y_2\end{bmatrix}\right)\\ \Phi\left(\lambda\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) &= \lambda x_1 + \lambda i x_2 = \lambda(x_1 + ix_2) = \lambda\Phi\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right).\end{aligned}$$

***Special mappings**:* *<u>Injective, Surjective, Bijective</u>*

Definition 2.16 (Injective, Surjective, Bijective)**.** Consider a mapping $\Phi : \mathcal{V} \to \mathcal{W}$, where $\mathcal{V}, \mathcal{W}$ can be arbitrary sets. Then Φ is called

- *Injective* if $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{V} : \Phi(\boldsymbol{x}) = \Phi(\boldsymbol{y}) \implies \boldsymbol{x} = \boldsymbol{y}$.
- *Surjective* if $\Phi(\mathcal{V}) = \mathcal{W}$.
- *Bijective* if it is injective and surjective.



The definition and figure examples of these terms.

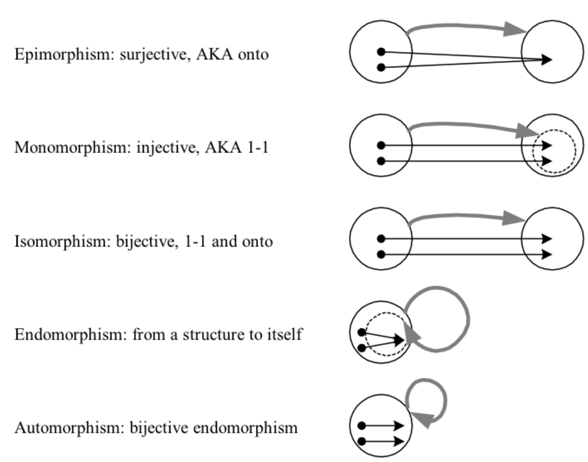
Special cases of linear mappings:

Definitions

- *Isomorphism:* $\Phi : V \to W$ linear and bijective
- *Endomorphism:* $\Phi : V \to V$ linear
- *Automorphism:* $\Phi : V \to V$ linear and bijective
- We define $\mathrm{id}_V : V \to V$, $\boldsymbol{x} \mapsto \boldsymbol{x}$ as the *identity mapping* or *identity automorphism* in V.

- **Epimorphism**: a homomorphism that is surjective (AKA onto)
- **Monomorphism**: a homomorphism that is injective (AKA one-to-one, 1-1, or univalent)
- **Isomorphism**: a homomorphism that is bijective (AKA 1-1 and onto); isomorphic objects are equivalent, but perhaps defined in different ways

Graph understanding



- **Automorphism**: a bijective endomorphism (an isomorphism from an object onto itself,

- **Endomorphism**: a homomorphism from an object to itself

essentially just a re-labeling of elements)

For more…

Homomorphism

In algebra, a homomorphism is a structure-preserving map between two algebraic structures of the same type (such as two groups, two rings, or two vector spaces). The word homomorphism comes from the Ancient Greek language: ὁμός (homos) meaning "same" and μορφή (morphe) meaning "form" or "shape". However, the word was apparently

https://en.wikipedia.org/wiki/Homomorphism

Isomorphism

> Theorem
>
> Finite-dimensional vector spaces V and W are *isomorphism* if and only if $dim(V) = dim(W)$.

This theorem states that

- *there exists a linear, bijective mapping between two vector spaces of the same dimension, means they're kind of same thing*
- treat $\mathbb{R}^{m\times n}$ and $\mathbb{R}^{mn}$ the same

Properties of *isomorphism*

- For linear mappings $\Phi : V \to W$ and $\Psi : W \to X$, the mapping $\Psi \circ \Phi : V \to X$ is also linear.
- If $\Phi : V \to W$ is an isomorphism, then $\Phi^{-1} : W \to V$ is an isomorphism, too.
- If $\Phi : V \to W$, $\Psi : V \to W$ are linear, then $\Phi + \Psi$ and $\lambda\Phi$, $\lambda \in \mathbb{R}$, are linear, too.

Matrix Representation of Linear Mappings

Any n-dimensional vector space V is isomorphic to $\mathbb{R}^n$. Consider the ordered basis vectors of V, write as $B = (b_1, \ldots, b_n)$.

$B = (b_1, \ldots, b_n)$: ordered basis
$\{b_1, \ldots, b_n\}$: unordered basis
$[b_1, \ldots, b_n]$: a matrix whose columns are vectors $b_1, \ldots, b_n$

Coordinates

> *the coordinate vector/coordinate representation*
>
> **Definition 2.18** (Coordinates)**.** Consider a vector space V and an ordered basis $B = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ of V. For any $\boldsymbol{x} \in V$ we obtain a unique representation (linear combination)
>
> $$\boldsymbol{x} = \alpha_1 \boldsymbol{b}_1 + \ldots + \alpha_n \boldsymbol{b}_n \tag{2.90}$$
>
> of $\boldsymbol{x}$ with respect to B. Then $\alpha_1, \ldots, \alpha_n$ are the *coordinates* of $\boldsymbol{x}$ with respect to B, and the vector
>
> $$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbb{R}^n \tag{2.91}$$
>
> is the *coordinate vector/coordinate representation* of $\boldsymbol{x}$ with respect to the ordered basis B.

Transformation Matrix

> Definition

Definition 2.19 (Transformation Matrix). Consider vector spaces V, W with corresponding (ordered) bases $B = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ and $C = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_m)$. Moreover, we consider a linear mapping $\Phi : V \to W$. For $j \in \{1, \ldots, n\}$,

$$\Phi(\boldsymbol{b}_j) = \alpha_{1j}\boldsymbol{c}_1 + \cdots + \alpha_{mj}\boldsymbol{c}_m = \sum_{i=1}^{m} \alpha_{ij}\boldsymbol{c}_i \tag{2.92}$$

is the unique representation of $\Phi(\boldsymbol{b}_j)$ with respect to C. Then, we call the $m \times n$-matrix $\boldsymbol{A}_\Phi$, whose elements are given by

$$A_\Phi(i, j) = \alpha_{ij}\,, \tag{2.93}$$

the *transformation matrix* of Φ (with respect to the ordered bases B of V and C of W).

Example:

Algebra

Example 2.21 (Transformation Matrix)
Consider a homomorphism $\Phi : V \to W$ and ordered bases $B = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_3)$ of V and $C = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_4)$ of W. With

$$\begin{aligned} \Phi(\boldsymbol{b}_1) &= \boldsymbol{c}_1 - \boldsymbol{c}_2 + 3\boldsymbol{c}_3 - \boldsymbol{c}_4 \\ \Phi(\boldsymbol{b}_2) &= 2\boldsymbol{c}_1 + \boldsymbol{c}_2 + 7\boldsymbol{c}_3 + 2\boldsymbol{c}_4 \\ \Phi(\boldsymbol{b}_3) &= 3\boldsymbol{c}_2 + \boldsymbol{c}_3 + 4\boldsymbol{c}_4 \end{aligned} \tag{2.95}$$

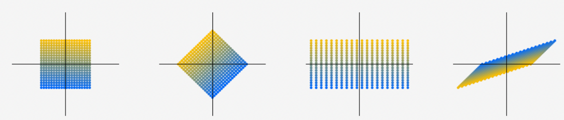
the transformation matrix $\boldsymbol{A}_\Phi$ with respect to B and C satisfies $\Phi(\boldsymbol{b}_k) = \sum_{i=1}^{4} \alpha_{ik}\boldsymbol{c}_i$ for $k = 1, \ldots, 3$ and is given as

$$\boldsymbol{A}_\Phi = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3] = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{bmatrix}, \tag{2.96}$$

where the $\boldsymbol{\alpha}_j,\ j = 1, 2, 3,$ are the coordinate vectors of $\Phi(\boldsymbol{b}_j)$ with respect to C.

Graphic

Example 2.22 (Linear Transformations of Vectors)



(a) Original data. (b) Rotation by 45°. (c) Stretch along the horizontal axis. (d) General linear mapping.

We consider three linear transformations of a set of vectors in $\mathbb{R}^2$ with the transformation matrices

$$\boldsymbol{A}_1 = \begin{bmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix}, \quad \boldsymbol{A}_2 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{A}_3 = \frac{1}{2}\begin{bmatrix} 3 & -1 \\ 1 & -1 \end{bmatrix}. \tag{2.97}$$

Basis Change

The change of the basis make the transformation matrix of a linear mapping replace by an equivalent matrix.

Theorem 2.20 (Basis Change). *For a linear mapping* $\Phi : V \to W$*, ordered bases*

$$B = (\boldsymbol{b}_1, \dots, \boldsymbol{b}_n), \quad \tilde{B} = (\tilde{\boldsymbol{b}}_1, \dots, \tilde{\boldsymbol{b}}_n) \tag{2.103}$$

of V *and*

$$C = (\boldsymbol{c}_1, \dots, \boldsymbol{c}_m), \quad \tilde{C} = (\tilde{\boldsymbol{c}}_1, \dots, \tilde{\boldsymbol{c}}_m) \tag{2.104}$$

of W*, and a transformation matrix* $\boldsymbol{A}_\Phi$ *of* Φ *with respect to* B *and* C*, the corresponding transformation matrix* $\tilde{\boldsymbol{A}}_\Phi$ *with respect to the bases* $\tilde{B}$ *and* $\tilde{C}$ *is given as*

$$\tilde{\boldsymbol{A}}_\Phi = \boldsymbol{T}^{-1} \boldsymbol{A}_\Phi \boldsymbol{S} \,. \tag{2.105}$$

Here, $\boldsymbol{S} \in \mathbb{R}^{n \times n}$ *is the transformation matrix of* id_V *that maps coordinates with respect to* $\tilde{B}$ *onto coordinates with respect to* B*, and* $\boldsymbol{T} \in \mathbb{R}^{m \times m}$ *is the transformation matrix of* id_W *that maps coordinates with respect to* $\tilde{C}$ *onto coordinates with respect to* C*.*

And we can have the *equivalence* and *similarity* defined by the Theorem 2.20.

Definition 2.21 (Equivalence). Two matrices $\boldsymbol{A}, \tilde{\boldsymbol{A}} \in \mathbb{R}^{m \times n}$ are *equivalent* if there exist regular matrices $\boldsymbol{S} \in \mathbb{R}^{n \times n}$ and $\boldsymbol{T} \in \mathbb{R}^{m \times m}$, such that $\tilde{\boldsymbol{A}} = \boldsymbol{T}^{-1} \boldsymbol{A} \boldsymbol{S}$.

Definition 2.22 (Similarity). Two matrices $\boldsymbol{A}, \tilde{\boldsymbol{A}} \in \mathbb{R}^{n \times n}$ are *similar* if there exists a regular matrix $\boldsymbol{S} \in \mathbb{R}^{n \times n}$ with $\tilde{\boldsymbol{A}} = \boldsymbol{S}^{-1} \boldsymbol{A} \boldsymbol{S}$

▼ **An basis change example**

Example 2.24 (Basis Change)
Consider a linear mapping $\Phi : \mathbb{R}^3 \to \mathbb{R}^4$ whose transformation matrix is

$$\boldsymbol{A}_\Phi = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{bmatrix} \tag{2.117}$$

with respect to the standard bases

$$B = (\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}), \quad C = (\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}). \tag{2.118}$$

We seek the transformation matrix $\tilde{\boldsymbol{A}}_\Phi$ of Φ with respect to the new bases

$$\tilde{B} = (\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}) \in \mathbb{R}^3, \quad \tilde{C} = (\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}). \tag{2.119}$$

Then,

$$\boldsymbol{S} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \qquad \boldsymbol{T} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{2.120}$$

where the ith column of $\boldsymbol{S}$ is the coordinate representation of $\tilde{\boldsymbol{b}}_i$ in terms of the basis vectors of B. Since B is the standard basis, the coordinate representation is straightforward to find. For a general basis B, we would need to solve a linear equation system to find the λ_i such that

$\sum_{i=1}^{3} \lambda_i \boldsymbol{b}_i = \tilde{\boldsymbol{b}}_j$, $j = 1, \ldots, 3$. Similarly, the jth column of $\boldsymbol{T}$ is the coordinate representation of $\tilde{\boldsymbol{c}}_j$ in terms of the basis vectors of C.

Therefore, we obtain

$$\tilde{\boldsymbol{A}}_\Phi = \boldsymbol{T}^{-1} \boldsymbol{A}_\Phi \boldsymbol{S} = \frac{1}{2} \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 \\ 0 & 4 & 2 \\ 10 & 8 & 4 \\ 1 & 6 & 3 \end{bmatrix} \tag{2.121a}$$

$$= \begin{bmatrix} -4 & -4 & -2 \\ 6 & 0 & 0 \\ 4 & 8 & 4 \\ 1 & 6 & 3 \end{bmatrix} . \tag{2.121b}$$

Image and Kernel

Definition

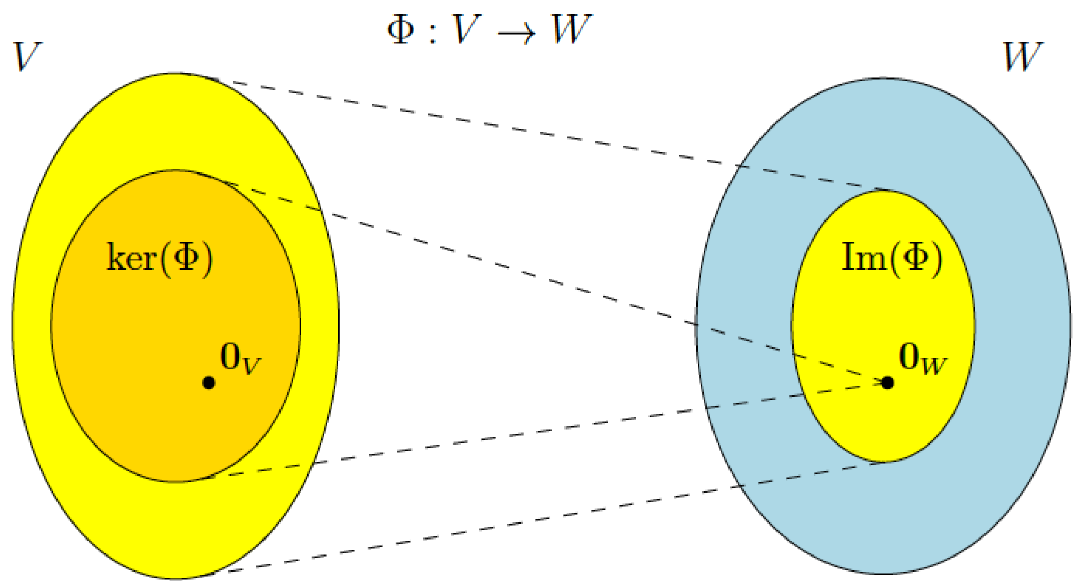
Definition 2.23 (Image and Kernel)**.**
For $\Phi : V \to W$, we define the *kernel/null space*

$$\ker(\Phi) := \Phi^{-1}(\boldsymbol{0}_W) = \{\boldsymbol{v} \in V : \Phi(\boldsymbol{v}) = \boldsymbol{0}_W\} \tag{2.122}$$

and the *image/range*

$$\operatorname{Im}(\Phi) := \Phi(V) = \{\boldsymbol{w} \in W | \exists \boldsymbol{v} \in V : \Phi(\boldsymbol{v}) = \boldsymbol{w}\} . \tag{2.123}$$

We also call V and W also the *domain* and *codomain* of Φ, respectively.



The kernel is the set of vectors $v \in V$ that maps onto the neutral element $0_w \in W$.

The image is the set of vectors $w \in W$ that can be reached by mapping from any vector in V.

Remark. Consider a linear mapping $\phi : V \to W$, where V, W are vector spaces.

- The null space is never empty, since $\phi(0_v) = 0_w$, and $0_v \in \ker(\phi)$
- $\mathrm{Im}(\phi)$ is a subspace of W, $\ker(\phi)$ is a subspace of V.
- ϕ is injective (one-to-one) if and only if $\ker(\phi) = \{0\}$.

Remark. (Null space and Column space). Consider $A \in \mathbb{R}^{m\times n}$ and a linear mapping $\phi : \mathbb{R}^n \to \mathbb{R}^m, x \mapsto Ax$.

- The image is the span of the columns of A, also called column space. For $A = [a_1, \ldots, a_n]$, where a_i are the columns of A, we obtain

$$\mathrm{Im} = \{Ax : x \in \mathbb{R}^n\} = \left\{ \sum_{i=1}^{m} x_i a_i : x_1, \ldots, x_n \in \mathbb{R} \right\} = \mathrm{span}[a_1, \ldots, a_n] \subseteq \mathbb{R}^{m\times n}$$

- $\mathrm{rk}(A) = \dim(\mathrm{Im}(\phi))$ = number of pivots = number of linear independent rows
- The kernel/null space $\ker(\phi)$ is the general solution to the homogeneous system of linear equation $Ax = 0$ and captures all possible linear combinations of the elements in $\mathbb{R}^n$ that produce $0 \in \mathbb{R}^m$.

Affine Spaces (linear space)

The spaces that are offset from the origin.

Affine subspaces

> Definition
>
> **Definition 2.25** (Affine Subspace)**.** Let V be a vector space, $\boldsymbol{x}_0 \in V$ and $U \subseteq V$ a subspace. Then the subset
>
> $$L = \boldsymbol{x}_0 + U := \{\boldsymbol{x}_0 + \boldsymbol{u} : \boldsymbol{u} \in U\} \tag{2.130a}$$
> $$= \{\boldsymbol{v} \in V | \exists \boldsymbol{u} \in U : \boldsymbol{v} = \boldsymbol{x}_0 + \boldsymbol{u}\} \subseteq V \tag{2.130b}$$
>
> is called *affine subspace* or *linear manifold* of V.
>
> U is called *direction/direction space*, and x_0 is called *support point*.

Affine mapping

> For tow vector spaces V, W, a linear
>
> mapping $\Phi : V \to W$, and $\boldsymbol{a} \in W$, the mapping
>
> $$\phi : V \to W \tag{2.132}$$
> $$\boldsymbol{x} \mapsto \boldsymbol{a} + \Phi(\boldsymbol{x}) \tag{2.133}$$
>
> is an *affine mapping* from V to W. The vector $\boldsymbol{a}$ is called the *translation vector* of ϕ.
>
> - Every affine mapping $\phi : V \to W$ is also the composition of a linear mapping $\Phi : V \to W$ and a translation $\tau : W \to W$ in W, such that $\phi = \tau \circ \Phi$. The mappings Φ and τ are uniquely determined.
> - The composition $\phi' \circ \phi$ of affine mappings $\phi : V \to W$, $\phi' : W \to X$ is affine.
> - Affine mappings keep the geometric structure invariant. They also preserve the dimension and parallelism.